Deep Learning —

Biologist + Computer programmer
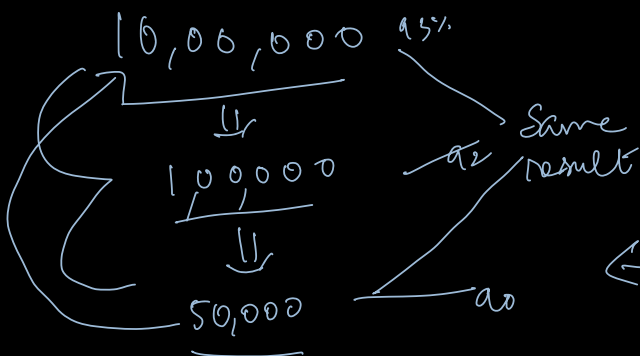
Neural Networks —

Machine Learning	Deep Learning
Columns ↑ increases ML struggle Curse of Dimensionality ←→	Exponential amount of Dimensionality

$10,00,000$ 95%
$\Downarrow$
100000 $\xrightarrow{q_2}$ Same result
$\Downarrow$
$50,000$ $\xleftarrow{a_0}$

humungous amount of Sample

←→

Simple task can be handled

1000×1000 Complex task
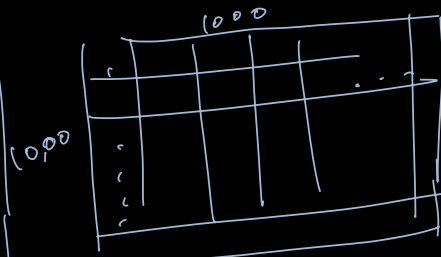
Maruti 800

$$\frac{1 \text{ litre} \rightarrow 20 \text{ km/ltr}}{\text{Speed} - 80 \text{ km/hr}}$$

Ferrari

$$\frac{1 \text{ litle} \rightarrow \ £5 \text{ km}}{\text{Speed} \rightarrow 250 \text{ km/hr}}$$

daily
Office

commial
traffic

Outing

Himalayas

Budget

Common Human

Neural
Networks

AI   DL   ML

1960

expert system

2010

Jeoffry Hinton

Neural Network

Neuron

Multiplication and addition

$$x \xrightarrow{\ w\ } \text{(O)} \xrightarrow{\ } Y$$

input   output
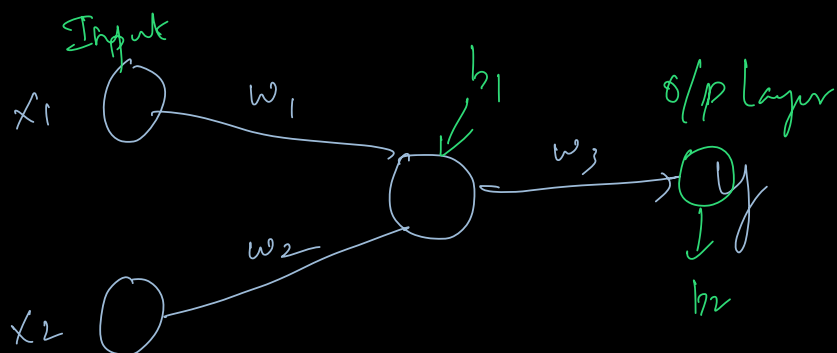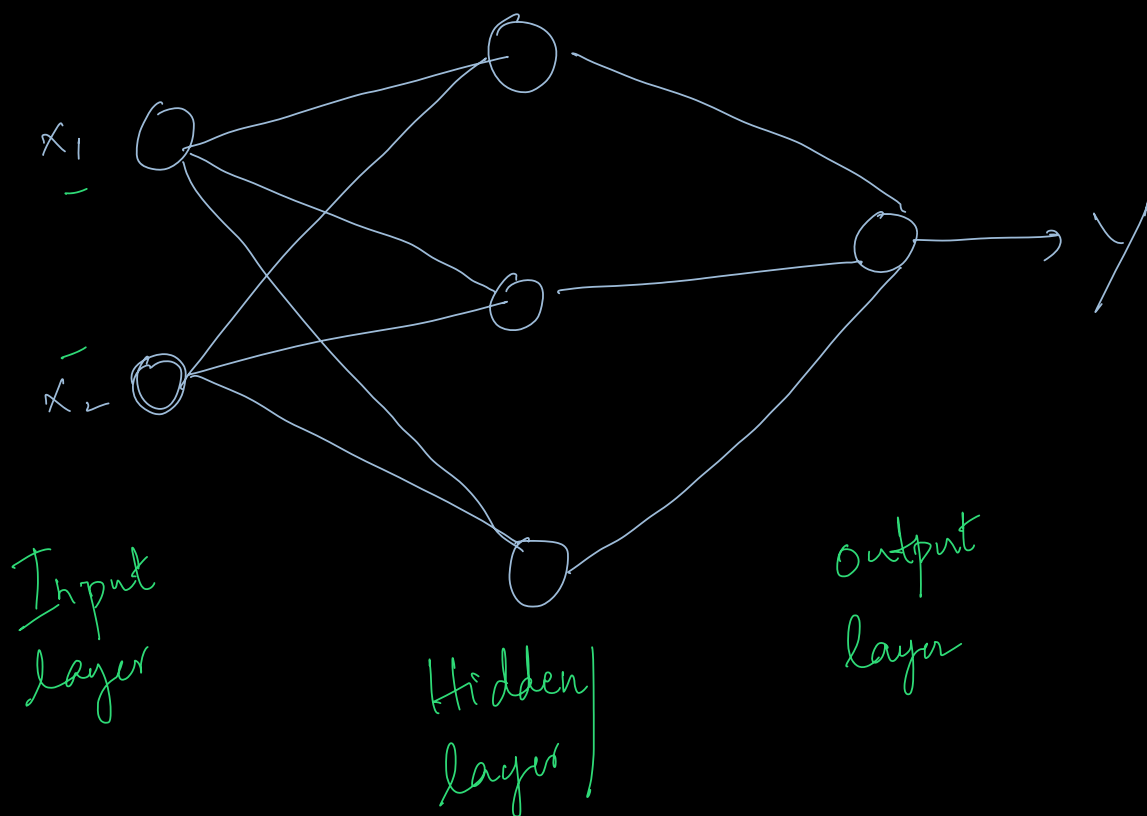
$$y = f(x)$$

$$f(x) = x \cdot w + b$$

$$Y = \left(\left(x_1 \cdot w + b_1\right) \cdot w_2 + b_2\right) \cdot w_2 + b_3$$

Input

x_1 w_1 b_1 o/p layer

         w_3 y

w_2

x_2 b_2

$$y = \left(\left(X_1 \cdot w_1 + b_1 \right) + \left(X_2 \, w_2 + b_1 \right) \right) \cdot w_3 + b_2$$

x_1

x_2 y

Input
layer   Hidden
            layer   Output
                                   layer

x_1 —

x_2 —

Input layer

Hidden 1 layer

Hidden 2 layer

Output layer

---

x_1

w_1

b

w_2

x_2

Y
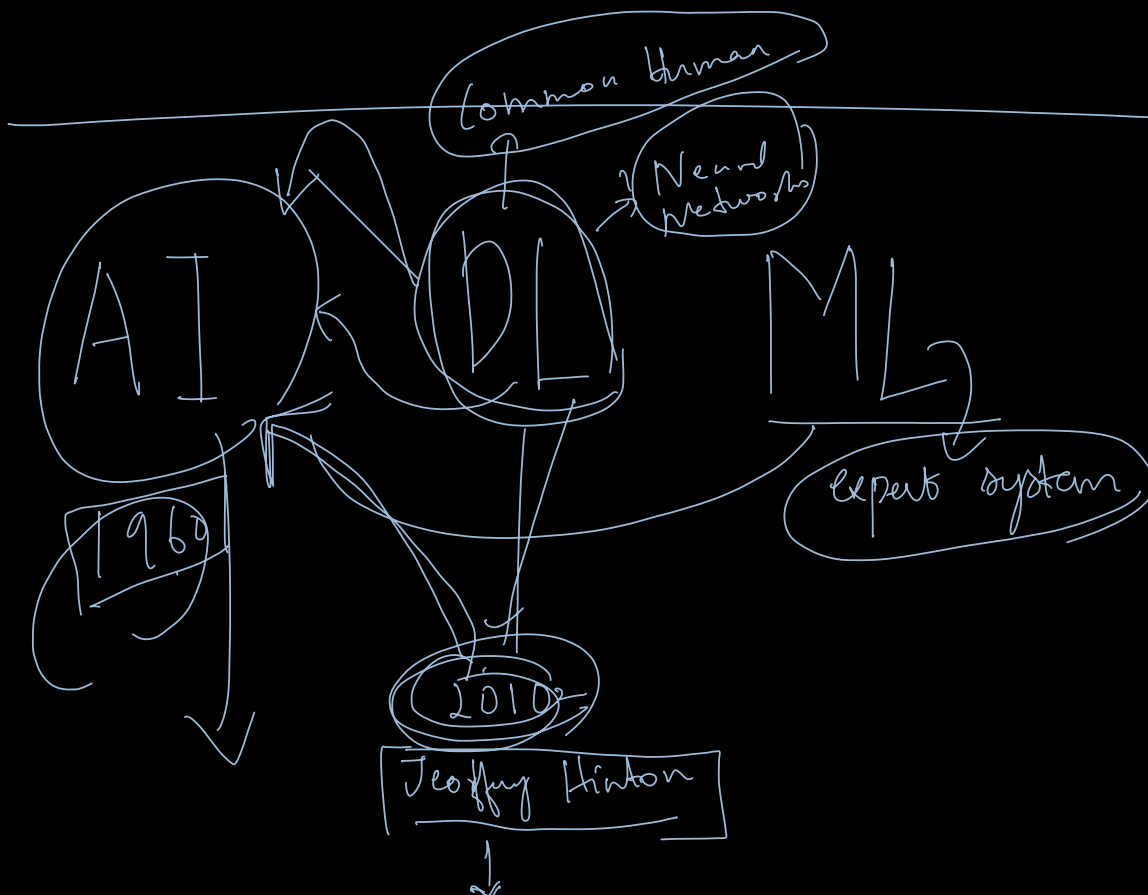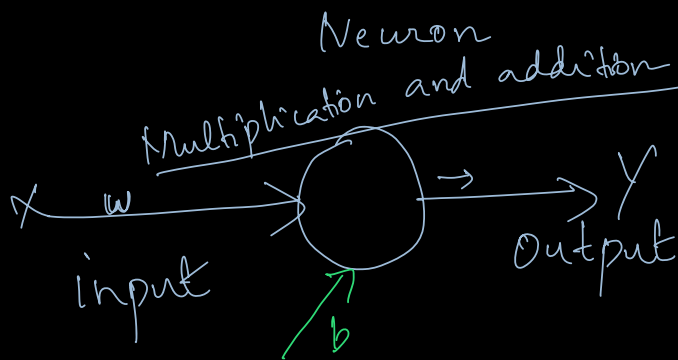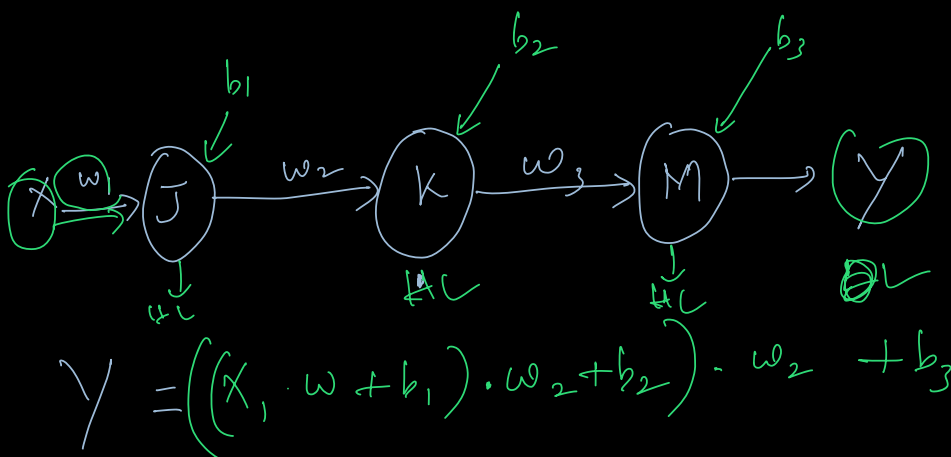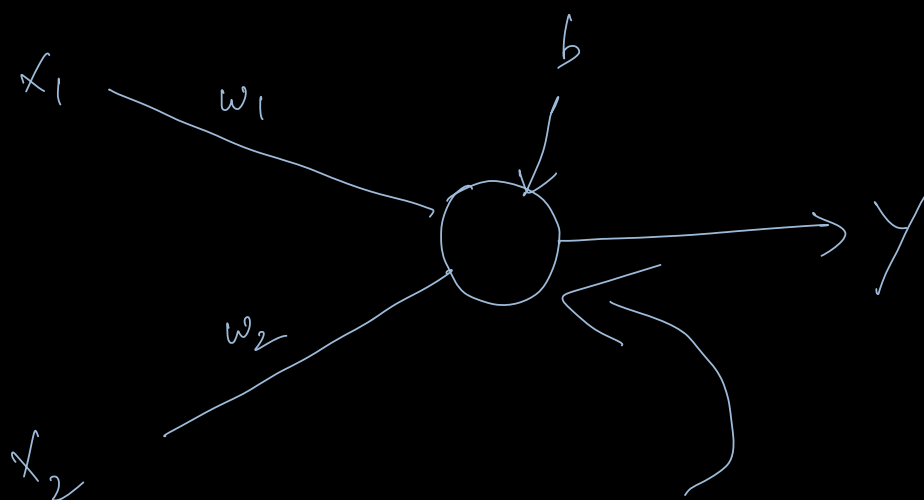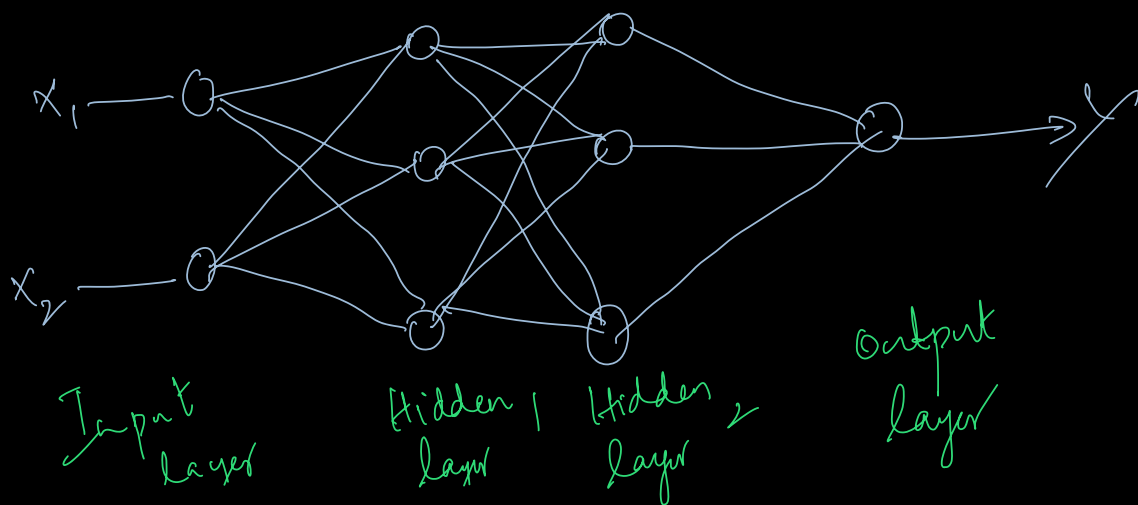
$$Y = x_1 W_1 + x_2 W_2 + b$$

Linear Regression

Linear Information

x ↑	y ↑
x ↑	y ↓
x ↓	y ↑
x ↓	y ↓

x_1 ———→

b

x_2 ———→ ○ → [ ] → y

Activation function
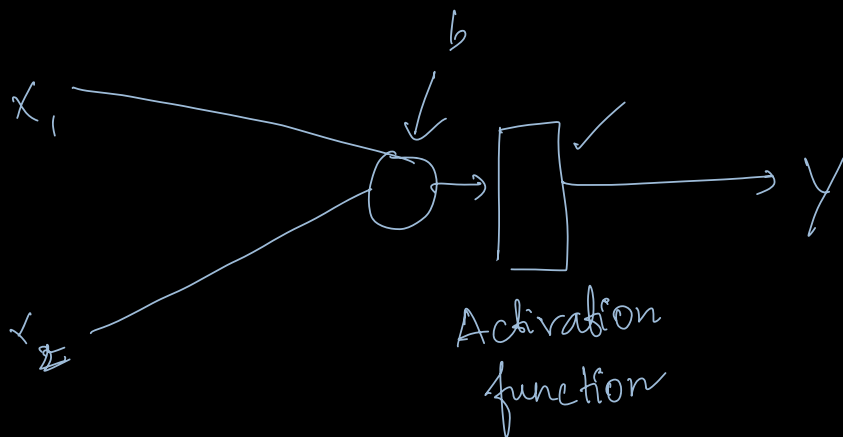
Activation function =
Identity function.

$5 \cdot I \Rightarrow 5$

$100 \cdot I \Rightarrow 100$

$$\begin{cases} \text{Identity function} \implies \text{Linear Regression} \checkmark \\ \text{Sigmoid function} \implies \text{Logistic Regression} \checkmark \end{cases}$$

$$\frac{1}{1 + e^{-(m x + b)}}$$

Linear Regression

10 Columns $\longrightarrow$ $\underline{11 \text{ parameters}}$

10 slope   1 Intercept

$$a, b \implies 1, a, b, a^2, b^2, 2ab +$$

$$2ab^2 +$$

feed forward layer

Back-Propagation

$$2ab$$

a, b
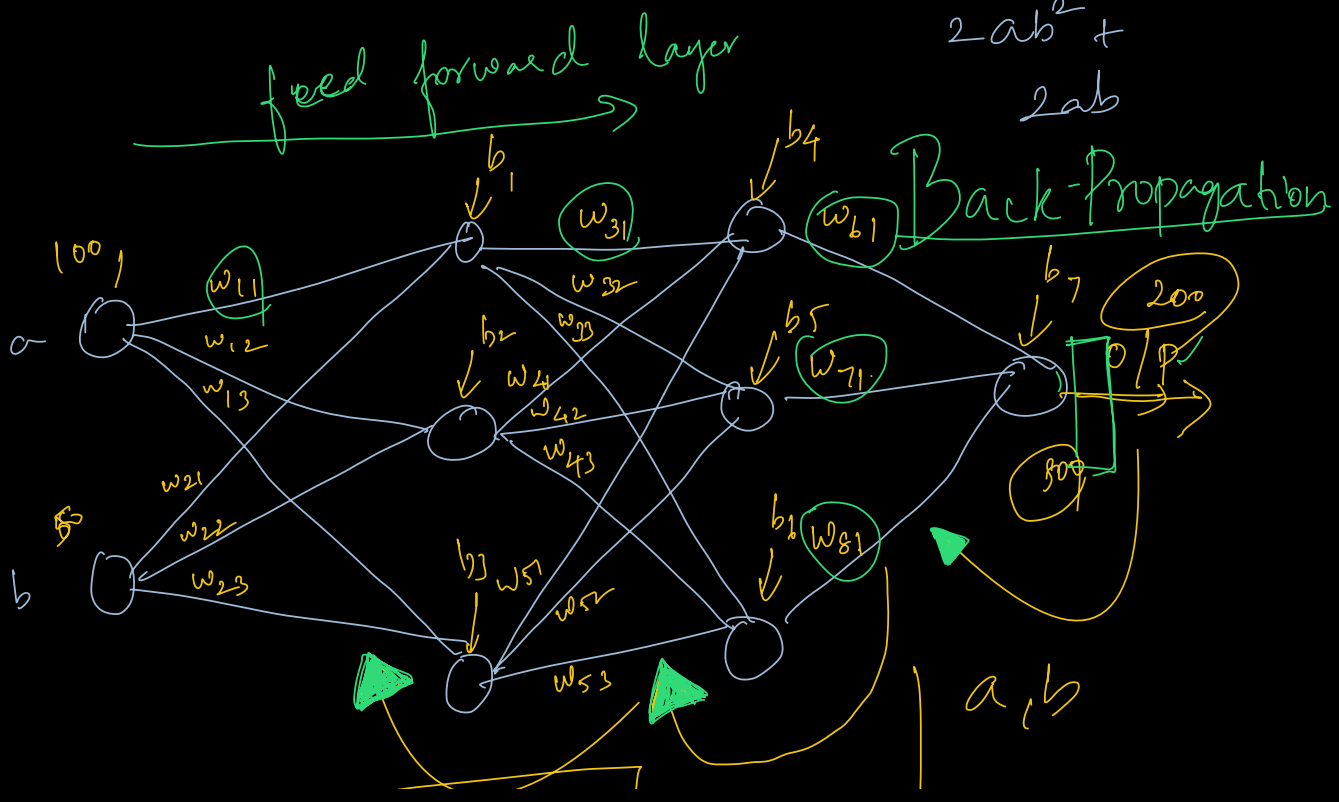
| 25 parameters |

$$2 \times 3 \times 3$$

$$18 + (3 + 3 + 1)$$

$$\underline{25} \checkmark$$

3 Parameters

$$\bar{ab}, a^2, b^2, 2ab$$

$$\underline{6} \rightarrow$$

2 Hidden Layers

10 Neurons

$$(2 \times 10 \times 10)$$

$$(200) + (10 + 10 + 1)$$

| 221 | $\rightarrow$ parameter

Deeply

task $\swarrow$ NLC $\checkmark$

| 5 people | $\rightarrow$ pm

task $\checkmark$ DL

| 500 people |

High quality
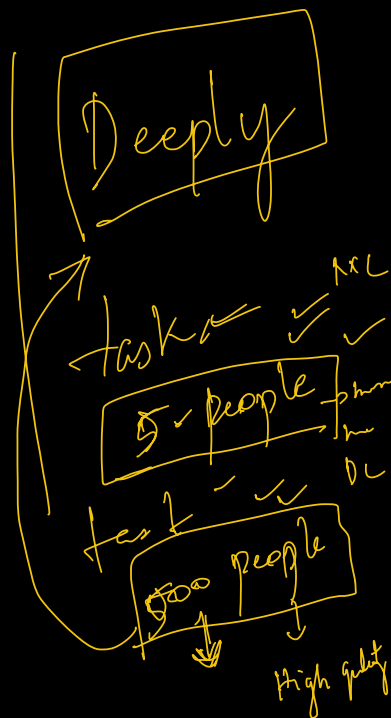
$$\left\{ \begin{array}{l} \text{Sigmoid function} \\ \text{Tanh function} \\ \text{Softmax function} \\ \text{Step function} \end{array} \right\} \rightarrow$$

1960s

before
2000s $\rightarrow$

$\downarrow$ | 2010

RELU function $\Big\}$ μ 2000 n+1 $\downarrow$

ELU function

Auto Encoders Unsupervised

B1
→ 10 → (W_n random) ⟶ Back propagation
W_n Optimised 1

B2
→ co ⟶ W_n Optimised 1 ⟶ Back propagation
W_n Optimised 2

B2
→ co ⟶ W_n Optimised 2 ⟶ Back
W_n Optimus 3

(Epoch Iteration)
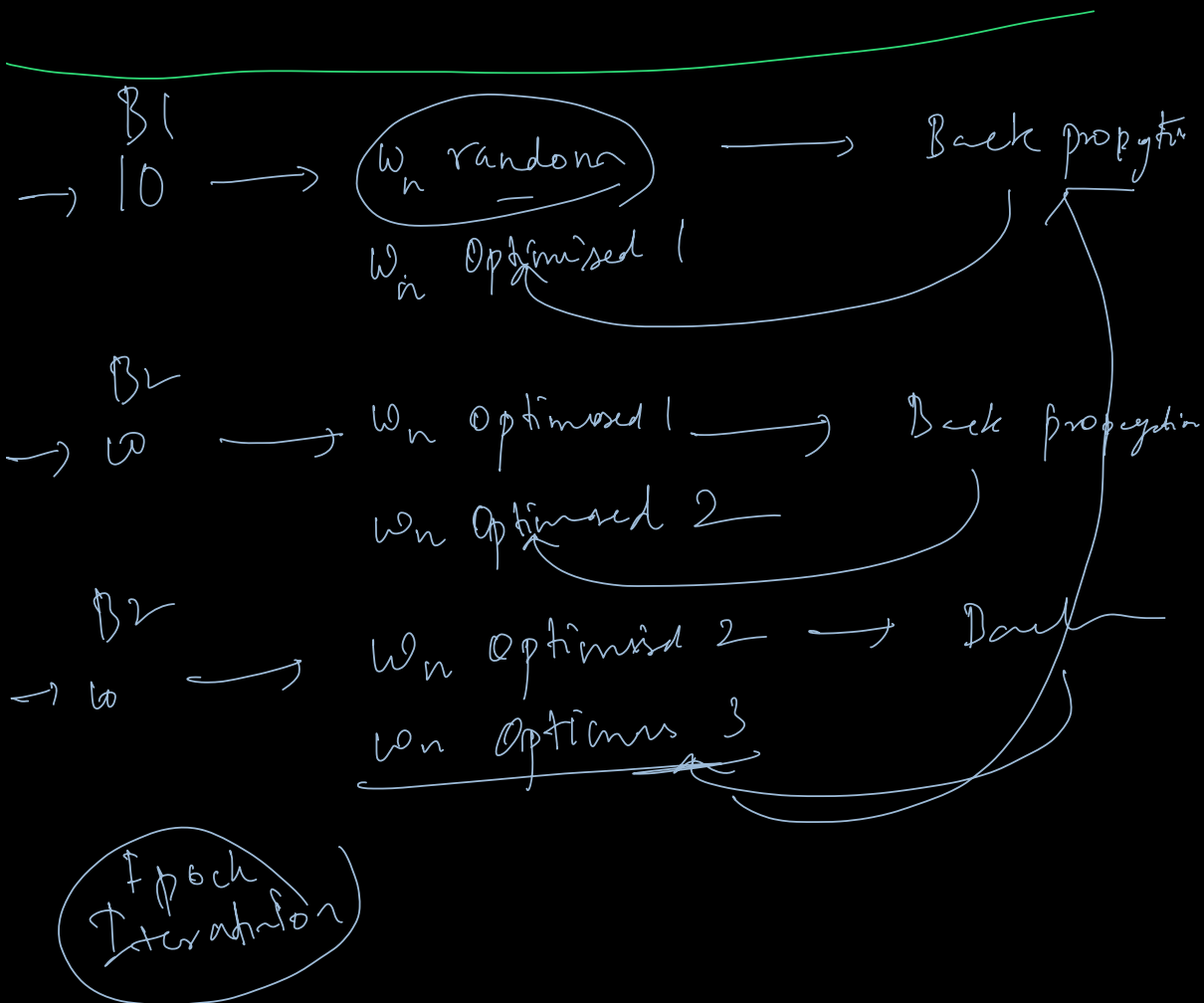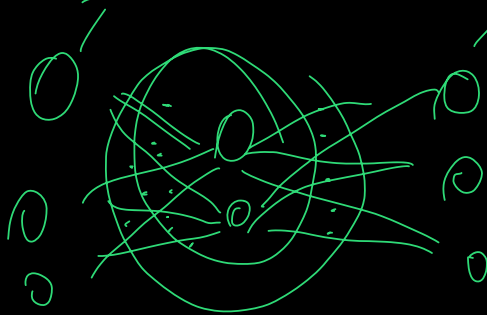
Why NN was ~~Abondoned~~

1. <u>Vanishing</u> Gradient descent
2. Exploding Gradient descent

---

1. Vanishing Gradient descent | Exploding god

$$m = \left[m - \left(\frac{6}{6m} \right) \cdot l_r \right] \longrightarrow ?$$

0

$$\boxed{m = \text{constant}}$$

→ min error